DETECTION AND IDENTIFICATION OF BACTERIA USING AUTOMATED BIOLOGICAL SAMPLE PROCESSING WITH A RELATIONAL DATABASE MANAGEMENT SYSTEM

Rabih E. Jabbour*
Geo-Centers, Incorporation
Gunpowder branch P.O. Box: 68
Aberdeen Proving Ground, MD 21010-0068
Fax: 410-436-3764
rjabbour@geo-centers.com

Samir V. Deshpande
Science & Technology Corporation
Edgewood, MD 21040-2734.
Samir.Deshpande@c-mail.apgea.army.mil

A. Peter Snyder
US Army Edgewood Chemical Biological Center
Soldier Biological and Chemical Command
Aberdeen Proving Ground, Maryland, 21010-5424
peter.snyder@us.army.mil

Abstract: The biological sample processing system (BSPS) combines sample pretreatment, protein concentration, and separation modules and was interfaced to electrospray ionization mass spectrometry (ESI-MS). Multivariate analysis software capabilities were used for automated data processing. The bacterial identification by the BSPS-MS is based on Liquid chromatography-mass spectrometry analysis of various bacterial proteins followed by their correlation with in-house database constructed from several bacteria of interest using our in-house ProMAPDB software. The identification of bacteria using ProMAPDB was validated using standard proteins analyzed by the BSPS-MS under various conditions. The results show that BSPS analysis was reproducible and sensitive at 1000 cfu/mL.

INTRODUCTION

Detection and identification of pathogens of biological origin are of great importance to the armed forces and civilian sector. The detection and identification of microorganisms may be efficiently achieved through the analysis of their protein profiles [1,2]. Protein constitutes greater than 50% of the dry weight of microorganism cellular components [3-7] and may be an ideal class of biomarker for bacterial

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
01 OCT 2005	**N/A**	**-**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Detection And Identification Of Bacteria Using Automated Biological Sample Processing With A Relational Database Management System	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Geo-Centers, Incorporation Gunpowder branch P.O. Box: 68 Aberdeen Proving Ground, MD 21010-0068**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES
See also ADM001851, Proceedings of the 2003 Joint Service Scientific Conference on Chemical & Biological Defense Research, 17-20 November 2003. , The original document contains color images.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**UU**	**15**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

identification. Advancements in mass spectrometry (MS) ionization and detection methods make it a promising analytical technique to differentiate biological substances [8-11].

ESI-MS interrogates ions that are generated directly at atmospheric pressure from a liquid matrix [13]. Accordingly, ESI-MS is easily interfaced with sample pretreatment and chromatographic modules [14]. Such a hybrid system offers better run-to-run reproducibility [15], capability of on-line biological sample processing, and increased sensitivity [10] with respect to MALDI-MS. Moreover, by interfacing chromatographic techniques such as liquid chromatography (LC) with ESI-MS, most matrix interferences including buffer and detergent adducts are reduced and/or eliminated [16]. Interference removal allows for better mass spectral profile reproducibility of bacterial extracts and results in a more accurate deconvolution of the mass spectra of bacterial proteins [17,18]. Although capillary LC-ESI-MS overcomes many limitations found in MALDI analysis of microorganisms, the chromatography alone cannot eliminate all the interferences from a bacterial extract prior to ESI-MS analysis. The inability of the LC to eliminate non-proteinaceous interferences often results in ionization suppression and S/N reduction during ESI-MS analysis of bacterial proteins. Hence, a sample pretreatment module is necessary to separate the interferential species that are not compatible with LC such as cellular debris, large particulates and non-proteinaceous components.

The Biological Sample Processing System (BSPS) is an in-house automated system designed to serve as a front end for ESI-MS interrogation of bacterial proteins [19,20]. The BSPS combines an off-line sample-processing module with a chromatographic module interfaced to an ESI-MS for microorganism differentiation. The design of the BSPS is based on a simple, rugged, and reproducible sample-processing scheme that couples LC and ESI-MS instrumentation to provide flexibility, sensitivity, and reproducibility.

The sample pretreatment module of the BSPS consists of various processes that are necessary for an effective and reproducible bacterial analysis by ESI-MS. In general, the sample pretreatment module consists of a sequence of steps. The bacterial sample is lysed to disrupt the major cell structures in order to release the proteins. Cell lysis is followed by the removal of cellular debris and particulates from the bacterial extract using a size exclusion filtration process. The filtrate is then subjected to sequential molecular weight filtration and extraction processes to isolate and preconcentrate bacterial proteins prior to their LC separation. The chromatographic module of the BSPS consists of a conventional hydrophobic capillary LC column, and the capillary LC module was used in either the isocratic or gradient separation mode.

2. EXPERIMENTAL METHODS

2.1 BSPS parameters

The sonication method was integrated into the BSPS. 20 μL of lysed bacterial sample was subjected to a Microcon©-3 kDa size exclusion membrane from Millipore (MA, USA) and has a molecular weight cut off (MWCO) of 3 kDa. After size exclusion filtration, the bacterial extracts were transferred onto a hydrophobic C18 or C8 protein micro-trap cartridge from Michrom (CA, USA). Bacterial proteins from the liquid phase are captured onto a small band of solid phase in the trap to provide preconcentration and purification of the bacterial proteins. The bacterial proteins were eluted in a 20-50 μL sample volume from the protein micro-trap cartridge into the LC module using a six-valve LC port transfer system. The reverse phase (RP) LC columns were hydrophobic silica based C18 or C8 and were purchased from Michrom (CA, USA) and Phenomenex (CA, USA). The RPLC system utilized either a gradient or an isocratic mode of separation. The aqueous phase was 95/5/0.1-1% H_2O/ACN/AA, FA, or TFA and the organic phase was 90/10/0.1-1% ACN/H_2O/ AA, FA, or TFA.

2.2 Automated deconvolution Algorithm analysis

The in-house deconvolution algorithm was set to determine the bacterial protein masses through the analysis of their raw BSPS-MS analysis file. The deconvolution parameters for the number of scans and deconvoluted masses were selected prior to the start of the deconvolution process. The generated mass list was automatically compared with the in-house database of biological agents. The mass list usually consists of a large number of deconvoluted masses. The software reduces the mass list to reflect the bacterial protein masses at the set of integrated peaks. The total identification process was completed within 5-15 minutes. The deconvolution process is interfaced with relational database management software to update the in-house database with the experimental bacterial protein masses. A comparison between the in-house and the commercially available deconvolution software is list in the following table:

In house Software	***Commercial LCQ Software***
Produces masses based on scoring scheme. It assigns the base peak with the max charge state possible. This charge state is calculated based on the m/z spacing with the nearest peak. This will greatly reduces the analysis time from hours to few minutes.	Produces a distribution of possible masses at every mass interval. This resulted in longer time of computation especially when wide range of masses is selected.
It will eliminate the deconvolution of peaks that are below the threshold level (user-defined). This will reduce the time of deconvolution of large number of spectra.	Generates artifact peaks that could interfere with the true peaks. This will lead to over-interpret the MS spectrum.
Charge state of a peak is determined based on all above threshold peaks in the m/z spectrum. This is advantageous in the analysis of a mixture.	Assigning charge states to the corresponding deconvoluted masses in a given scan is tedious because of the generation of large number of possible masses.
Deconvolutes MS spectrum of a mixture even when two or more species overlap in their m/z ratio.	Charge state of a peak is determined based on a single pair of user-chosen peaks.
The deconvolution process is automated and no user interaction is needed during deconvolution.	The deconvolution process is manual and user based.

3 RESULTS AND DISCUSSION

3.1. Analysis of E. coli lysate with the BSPS-MS.

Vegetative *E. coli* samples dissolved in 5% ACN/ 0.1% TFA at a concentration of 10^6cfu/mL were lysed by vortex, heat or sonication. 10^6 *E. coli* cells in one milliliter of buffer were lysed, and ten microliters were injected into the LC module of the BSPS. Therefore, the equivalent amount of cells actually introduced into the system was 10^4. The TIC plots obtained from the BSPS-MS analyses of the *E. coli* lysates are shown in Fig. 1. The TIC plot of sonicated *E. coli* had relatively the highest signal intensity response and the largest number of peaks compared to the thermally treated and vortexed samples.

Deconvolution of the mass spectra for each chromatographic peak in an LC spectrum provides more representative information on the effect of lysis on the protein profile than visual inspection of the TIC plots. The deconvolution of the mass spectra can establish the bacterial protein masses for each *E. coli* lysate. Deconvolution of the mass spectra for each peak in the TIC plots (Fig. 1) using an in-house automated deconvolution algorithm resulted in at least 180, 160, and 165 protein masses for the sonication, thermal, and vortex methods, respectively. It has been reported that lysis of *E. coli* cells could result in the release of thousands of proteins in a broad concentration range with a wide range of pI values [21,22]. Therefore, it is unlikely to completely separate such a large number of proteins using a single dimension RPLC approach. Sonicated *E. coli* lysate had the largest number of deconvoluted protein masses than that vortexed or heated. In addition, the relative intensity for the common protein masses was the highest for sonicated *E. coli* than that vortexed or heated. However, the common protein masses

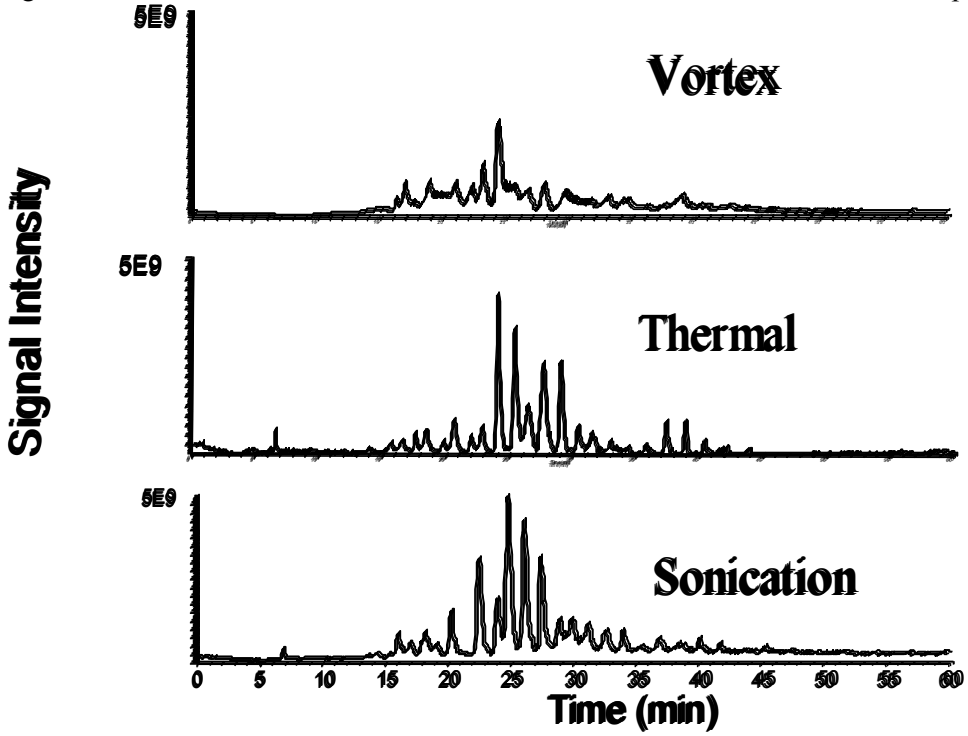

Figure 1. Comparison of the LC-MS TIC plots of *E. coli* protein extracts from different lysis methods. *E. coli* samples were dissolved in 5% ACN/0.1% TFA (v/v) and the protein extract injected was equivalent to 10^4 cells. Chromatographic conditions: Luna C8 column, 150 x 1 mm, flow rate was 30 µL/min, 90/10/0.1 H_2O/ACN/FA for A, 80/20/0.1 ACN/H_2O/FA for B, temperature 25 C.

deconvoluted from the thermally treated *E. coli* lysate had relatively the highest S/N ratios than that sonicated or vortexed. Fig. 2 shows a comparison between the mass spectra of a protein at 21.5 min. retention time from sonicated and thermally treated *E. coli* lysates. A relatively higher level of background noise was observed in the mass spectra of the sonicated *E. coli* than that heated. The increased background noise resulted from sonication being more effective than heating in the disruption of the cells and the release of more cellular components.

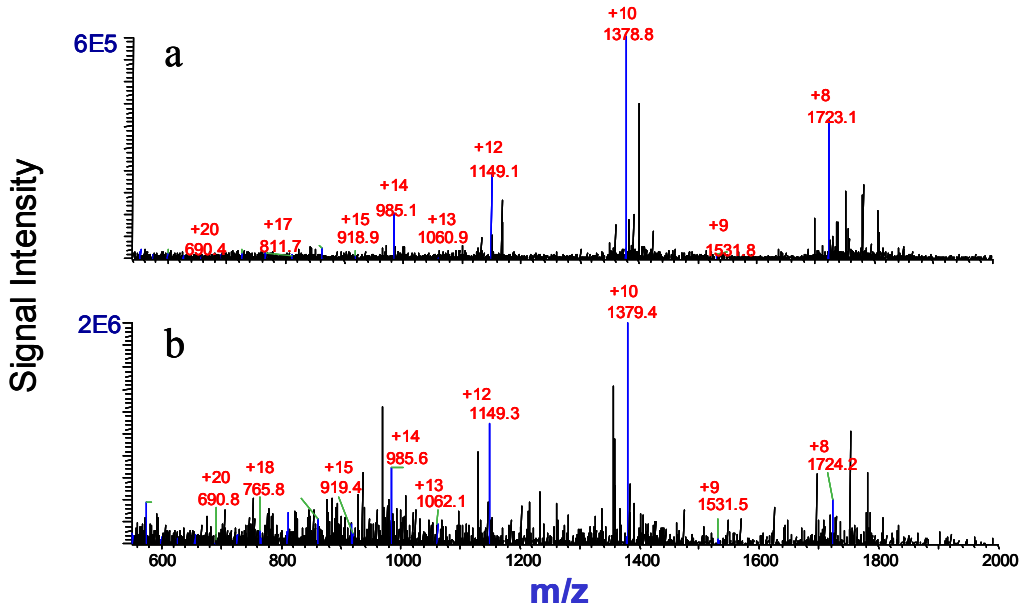

Figure 2. Comparison between the effect of (a) thermal lysis and (b) sonication on the mass spectra of *E. coli* protein extracts where the deconvoluted protein masses are 13,778.0 ± 0.7 Da and 13,784.1 ± 0.9 Da, respectively.

3.2. Analysis of BG lysates with the BSPS-MS

Sporulated BG samples dissolved in 5% ACN/ 0.1% TFA at a concentration of 10^6cfu/mL were lysed by vortex, heat and sonication. Samples containing 10^4 equivalent cells exposed to these lysis pretreatments were injected into the LC module of the BSPS. The TIC plots obtained from the BSPS-MS analyses of the BG lysates are shown in Fig. 3. The TIC plots were normalized relative to the plot with the highest signal intensity. The TIC plot of the sonicated BG had relatively the largest number of peaks and the highest signal intensity than that thermally treated or vortexed. The effect of the lysis method was more significant on the protein profile of BG spores (Fig. 3) than that of *E. coli* (Fig. 1). BG spores have sturdy cell walls (cortex) that require rigorous lysis in order to break open and release the largest possible amount of proteins. Sonication, and to a lesser extent heat, showed sufficient disruption of the BG spores as indicated by the protein profile observed in their TIC plots in Fig. 3.

The deconvolution of the mass spectra resulted in 8, 42, and 70 protein masses for the vortexed, heated, and sonicated BG lysates, respectively. Vortexed BG lysate had the least number of deconvoluted protein masses. The deconvoluted protein masses obtained from the vortexed BG were also found in the sonicated or heated BG preparation. However, mass spectra of the thermally treated BG lysate had relatively higher S/N ratios than that of sonication, and an example is shown in Fig. 4. The difference in the S/N ratios between the mass spectra of sonicated and heated BG lysates resembled that of *E. coli* (Fig. 2). The LC retention times of the protein are 15.2 and 20.5 minutes for Figs. 12a and b, respectively. Common proteins between heated and sonicated BG solutions displayed variations in their retention times that are attributed to the difference in the overall composition of the BG extracts. Thermal lysis of BG spores releases a relatively lower amount of biomolecules than that of sonication as observed in their corresponding TIC plots. On the contrary, the relatively greater number of competing molecules in the sonicated BG resulted in relatively shorter protein retention times than that of the thermally treated sample. This observation is in accordance with the solvophobic theory where proteins have a high retention factor on a nonpolar LC column when the solution contains neat aqueous eluents. The protein

retention value decreases as the content of the competing miscible eluent increases [42]. On the other hand, proteins in the thermally treated and sonicated *E. coli* extracts did not exhibit a significant change in their retention times. *E. coli* has a weaker cell wall than BG spores, therefore, lysis of the former by either the thermal or the sonication procedure did not result in a significant temporal change in the LC chromatogram protein extract pattern.

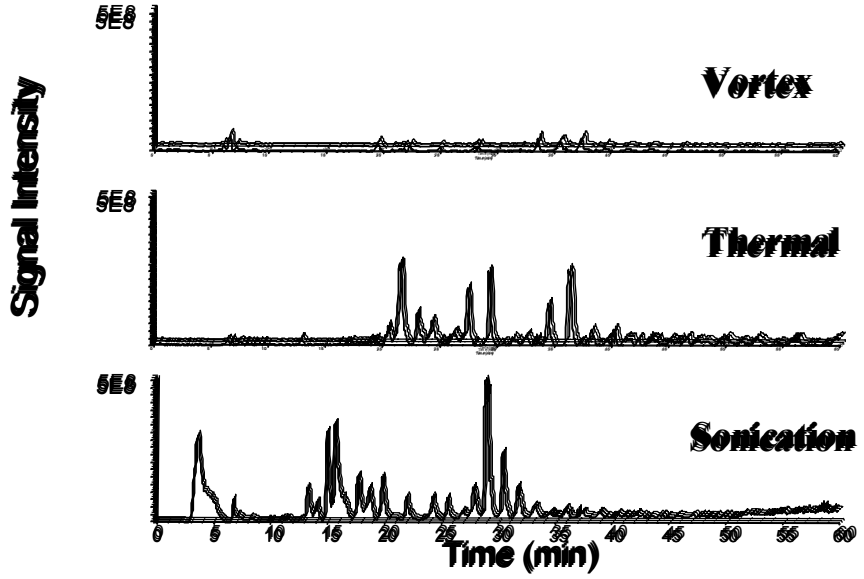

Figure 3. Comparison of the LC-MS TIC plots of BG protein extracts from different lysis methods. BG samples were dissolved in 5% ACN/0.1% TFA (v/v) and the injected protein extract was equivalent to 10^4 cells. Chromatographic conditions: Luna C8 column, 150 x 1 mm; flow rate was 30 µL/min; 90/10/0.1 H_2O/ACN/FA for A, 80/20/0.1 ACN/H_2O/FA for B; temperature 25 C.

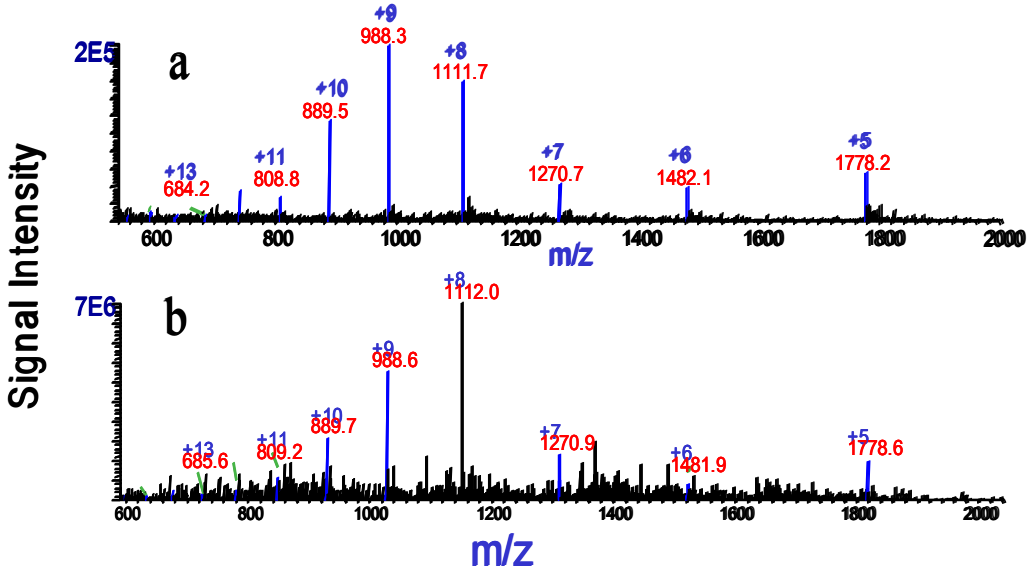

Figure 4. Comparison between the effect of (a) thermal lysis and (b) sonication on the mass spectra of BG protein extracts where the deconvoluted protein masses are 8,885.6 ± 0.5 Da and 8,888.1 ± 0.3 Da, respectively.

3.3. Reproducibility of the BSPS-MS analysis

The protein TIC pattern differences of the *E. coli* (Fig. 1) and BG (Fig. 3) extracts are essentially due to the lysis approach. However, the deconvoluted protein mass output, rather than a visual TIC plot comparison, determines the reproducibility of the BSPS-MS analyses of a particular bacterial extract under a given set of

conditions. Triplicate samples of *E. coli* and BG were prepared from the same respective batch. *E. coli*

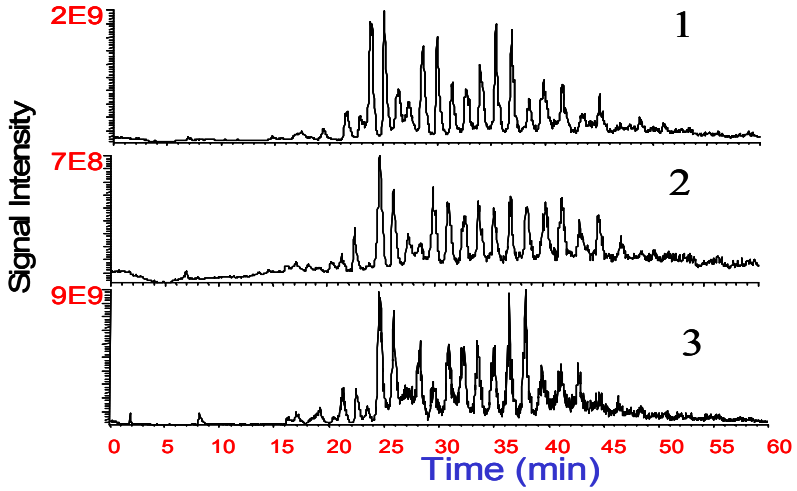

Figure 5. Replicate TIC plots for three separate bacterial protein extracts from sonicated *E. coli* samples. All three samples were collected from the same growth harvest and subjected to the same experimental conditions.

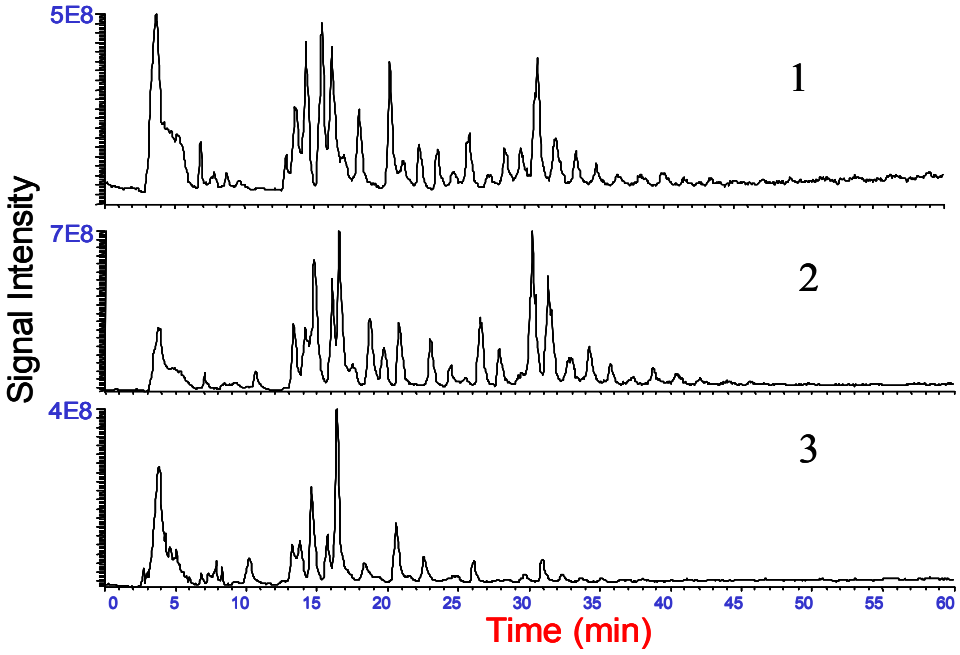

Figure 6. Replicate TIC plots from three separate bacterial protein extracts from sonicated BG samples. All three samples were collected from the same growth harvest and subjected to the same experimental conditions.

Table 1. Comparison of the most intense deconvoluted masses (S/N > 30) from E. coli protein extracts in Fig. 15. Reproducibility of 50%-72% was found upon comparing the mass lists (± 3 Da).

1	2	3
3522	**3521**	**3921**
3918	3589	3972
4496	**4495**	**4497**
5861	5438	5433
6415	5879	5853
7229	**6415**	**6413**
8534	**7739**	7211
11184	**8533**	**7736**
14236	8593	8538
16681	**11184**	9271
17582	**14239**	11178
18158	16687	**14241**
19296	**17581**	**18157**
24103	**18160**	19209
28468	**19296**	33288
35163	35167	**35163**
35201	**35204**	**35203**
40996	**40995**	**40994**

Table 2. Comparison of the most intense deconvoluted masses (S/N > 30) from sporulated BG protein extracts in Fig. 16. Reproducibility of 53%-73% was found upon comparing the mass lists.

1	2	3
5023	5027	5018
6254	**6255**	**6255**
7330	**6436**	**6437**
7334	**7335**	**7336**
8368	**8367**	**8367**
8889	**8887**	**8889**
9192	9188	**9195**
9532	**9530**	**9529**
10298	**10298**	**10298**
11183	**11185**	**11183**
13649	**13649**	13661
17513	**17511**	15657
20162	**19053**	**19053**
35163	19070	19066
35170	21120	20099
35176	**35162**	35523

and BG samples at concentrations of 10^6 cfu/mL were sonicated for 2 minutes, and ten microliters containing a protein extract equivalent to 10^4 cells were analyzed with the BSPS-MS. Figs. 5 and 6 show triplicate TIC plots of *E. coli* and BG, respectively. *E. coli* displays a relatively greater similarity in the TIC plots than that of BG. The difference in the signal intensity for certain peaks at the same retention time could be due to a combination of a co-elution of proteins and abundance variation. The number of deconvoluted protein masses were 180-182 and 69-74 for *E. coli* and BG, respectively. The number of common proteins among the triplicates was 141 and 59 for *E. coli* and BG, respectively. A comparison of the number of common proteins with that of the total number of deconvoluted proteins resulted in a 50%-73% reproducibility range for *E. coli* and BG. A portion of the total number of deconvoluted and similar masses for triplicate experiments of *E. coli* and BG are presented in Tables 1 and 2, respectively. These dominant protein masses were distinct, and a correlation of the common deconvoluted masses is observed to within ± 3 Da (masses in bold) regardless of the sample pretreatment conditions. It appears that the BSPS-MS analyses of the *E. coli* and BG samples exposed to the same experimental parameters were reasonably reproducible.

3.4. Sensitivity of the BSPS-MS analysis

E. coli samples harvested at 27 hours with different concentrations analyzed using the BSPS-MS. The TIC reveals a resemblance in the TIC pattern among the E. coli samples at various concentrations. This pattern similarity is observed at 20-50 minute, the time range where the maximum S/N ratios were observed. Moreover, the data showed a direct relationship between the concentration level and the signal

intensity as seen in their corresponding TIC signal intensities. The number of observed peaks for each

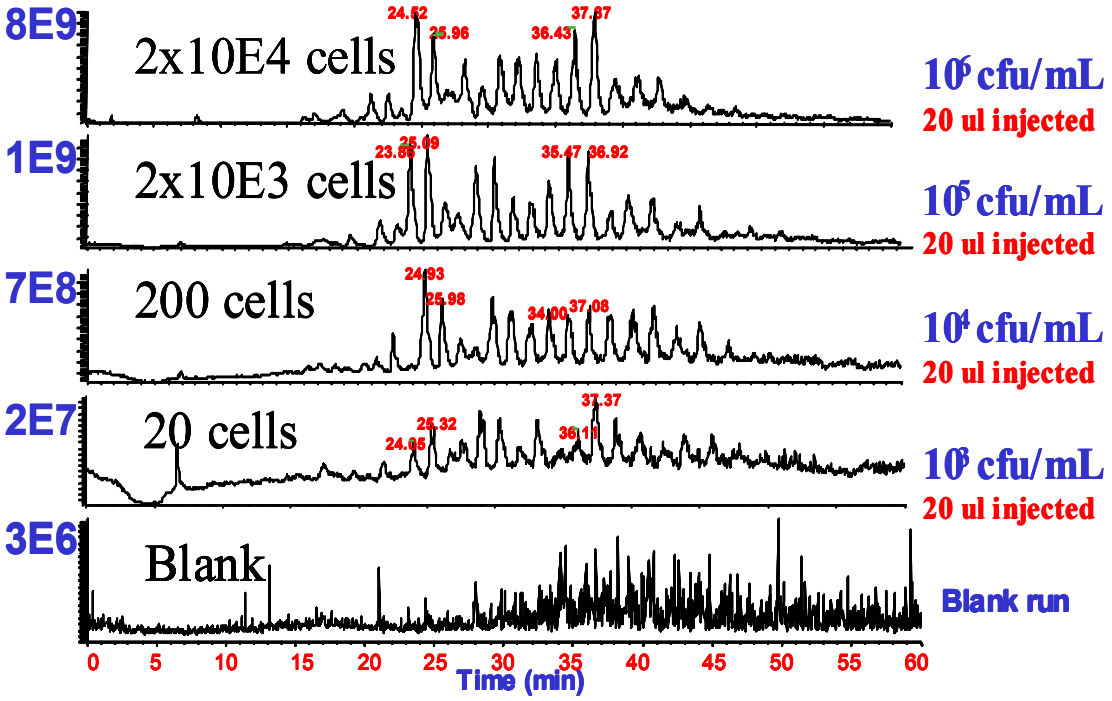

Figure 7: BSPS-MS analysis of E. coli samples harvested @ 27 hours with various concentration level

sample could not be interpreted to represent a single bacterial protein due to the fact that the bacterial extract contained large numbers of protein that are not individually resolved using single dimension LC.

Deconvolution of their corresponding mass spectra is an effective and a reliable approach to

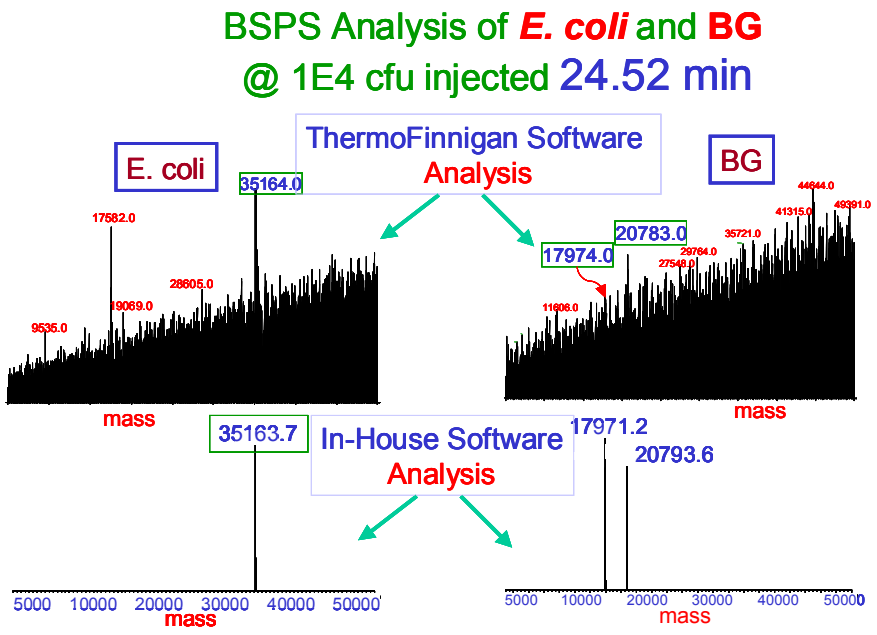

Figure 8: Comparison of the mass deconvolution output between the commercial and in-house algorithms. *E. coli* sample at concentration of 1E6 cfu/mL with 20uL sample volume injected onto the BSPS. (Top) represents the commercial masses deconvolution algorithm output at 24.52 minute, (bottom) represents the output of the in-house mass deconvolution algorithm at the same retention time.

elucidate the sensitivity of the BSPS to analyze E. coli samples presented a low concentration, figure 7. The sensitivity not only represents the TIC at low concentration but also the m/z information contained in each LC peak, which are used to determine the bacterial protein masses. The deconvolution could be achieved through selecting a time range in the TIC and determining manually, using the LCQ© software, the deconvoluted masses from their corresponding mass spectra. However, manual deconvolution is a lengthy process (5-7 hours for 60 minute BSPS-MS file) given the fact that hundreds of bacterial protein masses could be presented in a given BSPS-MS analysis. However, the application of an automated in-house deconvolution algorithm provided a rapid (5-10 minute for 60 minute BSPS-MS file) and accurate deconvolution masses for a given BSPS-MS raw file. In fact, a comparison of the deconvoluted masses derived from the in-house algorithm and the commercial one showed the presence of the same masses at a given time range. However, the commercial deconvolution algorithm is characterized with artifact peaks that are observed with the dominant masses of the bacterial proteins.

On the other hand, only the dominant masses in a given scan range are presented by the in-house algorithm. These advantages of the in-house deconvolution algorithm have great implication on developing an accurate and reliable experimental database that contains the dominant protein masses for bacterial identification. Figure .8 shows the fundamental difference between the commercial and in-house deconvolution algorithm. The top two graphs represent the output of the commercial deconvolution algorithm for *E. coli* and *Bacillus globigii* respectively. The bottom two graphs represent the output of the in-house deconvolution algorithm. The commercial algorithm provides artifact masses beside the true peaks (35164 for E. coli and 17974, 20783 for BG), which is not the case with the in-house algorithm. The user's set of parameters for every peak had to be selected manually to perform deconvolution process. The wider the deconvolution mass range, the greater the number of generated artifact masses. On the other hand, the in-house deconvolution algorithm is automated and the user's set parameters are once selected for the whole TIC prior to the deconvolution process, also the in-house deconvolution algorithm provided only the dominant masses regardless of the deconvolution mass range.

A comparison between the mass spectra for the solvent blank and the least concentrated *E. coli* indicated that the E. coli peaks observed at such low concentration have distinct charge state distributions that is not observed with the solvent blank spectra in the same time range. Upon deconvoluting their corresponding mass spectra to determine the possible masses of proteins in that time range a dominant mass of 9313 Da was observed for *E .coli* but not with the solvent blank mass spectra. The mentioned bacterial protein mass was reproducible and observed upon replicate analyses of *E. coli* samples at 20 equivalent cells. In fact, deconvolution of the solvent blank mass spectra showed no masses similar or with sufficient S/N to detect at that time range. It noteworthy to mention that a solvent blank run after the sample was analyzed did not provide any carry over from the previously analyzed *E. coli* sample. Accordingly, these experiments indicated that the BSPS has capability to reliably analyze *E. coli*, vegetative bacteria, with negligible carry over between consecutive BSPS-MS analyses.

3.5. Bacterial identification using the In-house relational Database software:

This work aimed at establishing an in-house database consists of the experimental proteome for different bacteria. The in-house database will be used to identify bacteria with high specificity due to its detailed structure and contents. This database had several layers divided according to specific parameters, such as sample preparation conditions of the bacteria, experimental parameters setup of the BSPS, etc. the software then perform the data processing followed by mining and sorting of the distinct and the uncommon masses in separate layers within the database. The common masses of a given bacteria obtained from repetitive runs of the bacterial lysates under certain experimental conditions are sorted in separated libraries. The distinct masses and the uncommon masses stored and updated periodically to monitor changes n the proteome of the bacteria. This specificity of the in-house database provides a high

confidence level of identification due to the presence of distinct masses extracted from repetitive runs of the bacteria under various experimental conditions. To investigate the impact of the in-house database on

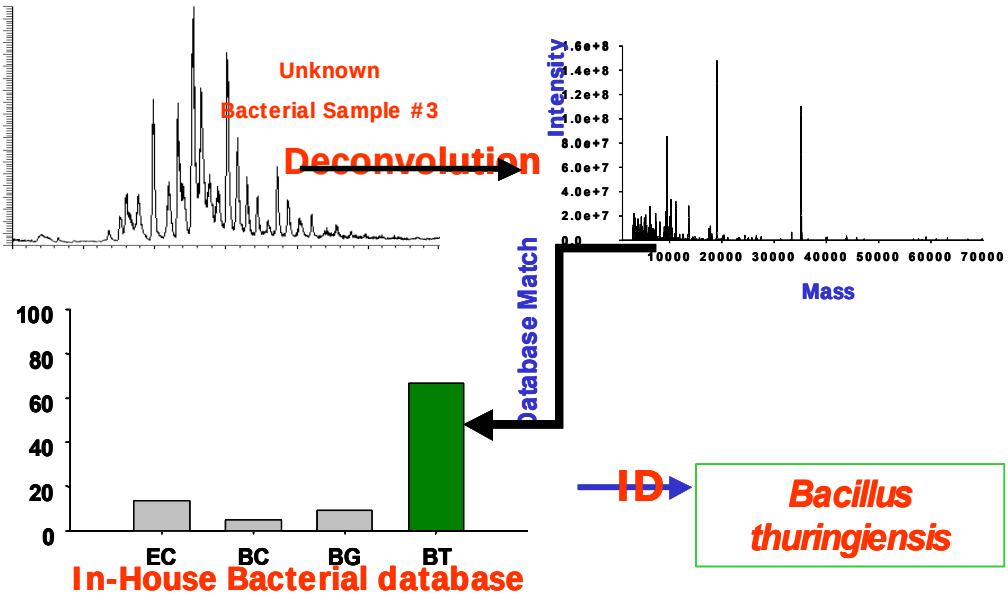

Figure 9: Identification of unknown bacterial lysate using the in-house relational database software (ProMAPDB). Unknown sample# 3, The TIC (upper left) was deconvoluted using in-house deconvolution algorithm and the mass list (upper right) was matched with the in-hose bacterial protein database.

the identification of the bacteria using the in-house relational database software, various unknown bacterial samples were analyzed using the BSPS-MS. The in-house automated deconvolution software was to process these samples. The resulted mass list was then matched against the in-house proteome database as shown in the figure 9. The result showed that this unknown sample has the highest match percentage with the experimental proteome masses of *bacillus thuringiensis*. The other three unknown samples were processed and showed highest matching percentage with their corresponding bacterial database.

Comparing the resulted data with the blind list of the unknown samples showed that all the unknown samples were identified correctly using the in-house relational database software. The utilization of in-house database consists of experimental protein masses is advantageous over the public databases, because the latter do not have equal number of experimental protein masses for different bacteria. for example, the E. coli proteome in the public database has at least 18000 entries in the Swiss-Prot and TrEMBL, only 646 entries are found for bacillus thuringiensis. Thus, identification of bacteria using the public database will provide statistically unrepresentative data of the actually identity of the bacteria once a given bacterial lysate, with minimum number of entries, is matched against all the bacterial proteome in the public database.

CONCLUSIONS

The automated deconvolution algorithm was effective in providing the bacterial protein masses. Such in-house system was verified with the commercial deconvolution software and the masses deconvoluted were highly similar with minute error range (2-5 ppm). The dominant masses are the only deconvoluted masses obtained using the in-house algorithm. This is important given the fact that to

establish a reliable database no artifact masses, often inherited by the commercial algorithm, should be eliminated.

The reproducibility of the deconvoluted masses provided a high degree of confidence in obtaining distinct masses necessary for differentiation of bacteria using the BSPS-MS. Correlation of the bacterial protein masses with the in-house database is scientifically sound than comparing with the public database. More investigations are needed to establish the in-house database as the reference library for future differentiation of more bacteria.

The sensitivity of the BSPS-MS analysis of bacteria was determined for vegetative and sporulated bacteria. The limit of detection was higher for vegetative bacteria than that of sporulated ones. The concentration of bacteria extract had an impact on the protein profile obtained by the BSPS-MS analysis. While there was number of bacterial protein masses that were observed in the most diluted samples the overall deconvoluted masses show a decrease in their correlation as concentration decreased. Proteins that are presented in large concentrations and amenable to efficient ionization were detected at the 20-100 equivalent cells level for vegetative and sporulated bacteria respectively.

REFERENCES

[1] Proceedings of the First Joint Services Workshop on Biological Mass Spectrometry, 1 (1997), Baltimore, Maryland.

[2] J.P. Anhalt, C. Fenselau, Anal. Chem. 47 (1975) 219.

[3] R.D. Holland, J.G. Wilkes, F. Rafii, J.B. Sutherland, C.C., Persons, K.J. Voorhees, J.O. Lay Jr., Rapid Commun. Mass Spectrom. 10 (1996) 1227.

[4] Y.P. Tan, Q. Lin, X.H. Wang, S. Joshi, C.L. Hew, K.Y. Leung, FEMS Infect. Immun. 215 (2002) 6475.

[5] O. Duché, F. Trémoulet, A. Namane, T.E.L.G. Consortium, J.C. Labadie, FEMS Microbiol. Lett. 215 (2002) 183.

[6] A.R. Hesketh, G. Chandra, A.D. Shaw, J.J. Rowland, D.B. Kell, M.J. Bibb, K.F. Chater, Mol. Microbiol. 70 (2002) 368.

[7] T.C. Cain, D.M. Lubman, W.J. Weber, Rapid Commun. Mass Spectrom. 10 (1996) 1026.

[8] M.J. Cole, C.G. Enke, Anal. Chem. 63 (1991) 1032.

[9] A. Fox, R.M. T. Rosario, L. Larsson, Appl. Environ. Microbiol. 59 (1993) 4354.

[10] C. Liu, S.A. Hofstadler, J.A. Bresson, H.R. Udseth, T. Tsukuda, R.D. Smith, A.P. Snyder, Anal. Chem. 70 (1998) 1797.

[11] S. Vaidyanathan, J.J. Rowland, D.B. Kell, R. Goodacre, Anal. Chem. 73 (2001) 4134.

[12] J.B. Fenn, M. Mann, C.K. Meng, S.F. Wong, C.M. Whitehouse, Mass Spectrom. Rev. 9 (1990) 37.

[13] F. Xiang, G.A. Anderson, T.D. Veenstra, M.S. Lipton, R.D. Smith, Anal. Chem. 72 (2000) 2475.

[14] T. Krishnamurthy, M.T. Davis, D.C. Stahl, T.D. Lee, Rapid Commun. Mass Spectrom. 13 (1999) 39.

[15] X. Liang, K. Zheng, M.G. Qian, D.M. Lubman, Rapid Commun. Mass Spectrom. 10 (1996) 1219.

[16] B.E. Chong, F. Yan, D.M. Lubman, F.R. Miller, Rapid Commun. Mass Spectrom. 15 (2001) 291.

[17] M.T. Kachman, H. Wang, D.R. Schwartz, K.R. Cho, D.M. Lubman, Anal. Chem. 74 (2002) 1779.

[18] R.D. Smith, J.A. Loo, C.G. Edmonds, C.J. Baringa, H.R. Udseth, Anal. Chem. 62 (1990) 882.

[19] A.R. Dongre, J. Eng, J.R. Yates III., Trends Biotechnol. 15 (1997) 383.

[20] M.C. Garcia *et al.,* J. Chromatogr. A 957 (2002) 187.

[21] http:www.expasy.ch

[22] C. Fenselau, in: C. Fenselau (Ed.) Mass Spectrometry for Characterization of Microorgansims, ACS Symposium Series, No. 541, American Chemical Society, Washington, DC, 1993.